

Exact solution of a p-spin model and its relationship to error correcting codes

Satish Babu Korada, Nicolas Macris
LTHC, I&C, EPFL, CH-1015
Email: {satish.korada, nicolas.macris}@epfl.ch

Abstract—An important quantity in the analysis of MAP decoding for LDPC codes is the conditional entropy of the input given the output. There exist conjectured formulas for this entropy derived from the replica technique and one sided bounds derived with the help of the interpolation method. In this paper we compute exactly such a quantity for a simpler spin model which retains the essential features of the communications problem. The result is a step towards a proof of the conjectured replica formula for the conditional entropy under MAP decoding.

I. INTRODUCTION

The close connection between coding theory and statistical physics uncovered by Sourlas [1] led to the application of tools such as the *replica technique* for the computation of noise thresholds in channel communication [2]. Broadly speaking error correcting codes can be viewed as spin models on graphs. There are two main classes of spin systems where the replica method is generally believed to yield exact formulas (for the phase transition thresholds, free energies, entropies ect...), namely models on complete graphs and models on random sparse graphs (or hypergraphs). The case of sparse graphs is particularly relevant to the coding community since the advent of good low density parity check (LDPC) codes.

The whole analysis based on the replica techniques is far from rigorous and relies on various uncontrolled assumptions. However in recent years there have been a number of advances concerning foundations of the replica formulas. These are based on the so-called *interpolation method* first developed by Guerra and Toninelli for the Sherrington-Kirkpatrick model [3]. Non trivial extensions to models on sparse graphs were then found by Franz and Leone [4] and in the context of LDPC and LDGM codes by Montanari [5]. All these works yield one sided bounds to the free energy (statistical physics) or the conditional entropy (coding). The bounds are believed to be tight but this is an unproven conjecture except for the case of the Sherrington-Kirkpatrick model for which Talagrand [6] proved that the celebrated Parisi solution is indeed exact.

In this contribution we consider a *p-spin model* on a complete hypergraph with a *gauge symmetry* built in the model. Here it suffices to say that this gauge symmetry is equivalent to channel symmetry for binary input memoryless channels, and that the model retains the main features of the channel communication problem. For this model we are able to prove that the replica formula for the free energy is exact.

Instead of defining directly the p-spin model we prefer to explain how the communication problem can be expressed in

statistical physics terms, and then show what modifications lead to the p-spin model. This we hope will make it clear that the present results are a step towards proofs of exactness of replica formulas for LDPC codes.

Consider communication over a binary input AWGN channel with transition probability $p(y|x)$ and noise variance h^{-1} . Here, we denote the input bits as x_i which take values in $\{-1, 1\}$ and the output as y_i . The bits x_i play the role of spins in statistical physics. Let $\underline{x} = \{x_1, \dots, x_n\}$ be the transmitted codeword from an LDPC code of rate $r = 1 - \frac{m}{n}$ where m is the number of check nodes, and $\underline{y} = \{y_1, \dots, y_n\}$ be the received sequence. It is easy to check that the half-log-likelihood ratios $h_i = \frac{1}{2} \ln \frac{p_{Y|X}(y_i|1)}{p_{Y|X}(y_i|-1)}$ have gaussian distribution with equal mean and variance $h_i \sim \mathcal{N}(h, h)$. Shannon's conditional entropy $H(\underline{X}|\underline{Y})$ is nothing else than the average Gibbs entropy of the posterior distribution $p_{\underline{X}|\underline{Y}}(\underline{x}|\underline{y})$ and is therefore directly related to the free energy of the spin model defined by the posterior distribution

$$H(\underline{X}|\underline{Y}) = \mathbb{E}_{\{h_i\}} [\ln Z] - nh$$

where the partition function of the spin model is

$$Z = \sum_{\underline{x} \in \{-1, +1\}^n} e^{\sum_{i=1}^n h_i x_i} \prod_{c=1}^m \frac{1}{2} (1 + x_{\partial c})$$

Here $x_{\partial c}$ denote the product $x_{i_1} \dots x_{i_k}$ where i_1, \dots, i_k are the variable nodes connected to the check node c .

The parity check constraints may be viewed as a limit of soft gaussian constraints,

$$\mathbb{E}_{\underline{h}} [\ln Z] = \lim_{J \rightarrow \infty} \mathbb{E}_{\{h_i, J_c\}} \left[\ln \sum_{\underline{x}} e^{\sum_{i=1}^n h_i x_i + \sum_{c=1}^m J_c (x_{\partial c} - 1)} \right]$$

with $J_c \sim \mathcal{N}(J, J)$. In this formula, the exponent is known as minus the hamiltonian. The spin model of interest is defined on a sparse Tanner graph and has a hamiltonian of the form $\mathcal{H}_{LDPC}(\underline{x}) = -\sum h_i x_i - \sum_{c=1}^m J_c x_{\partial c}$ with quenched couplings:

i) h_i are gaussians with equal mean and variance - a fact coming from channel symmetry

ii) J_c are gaussians with equal mean and variance (tending to ∞) - a convenient choice.

The distribution $p_{\underline{X}|\underline{Y}}(\underline{x}|\underline{y})$ is a Gibbs measure over $\{-1, +1\}^n$. The expectation with respect to this Gibbs measure is denoted by the bracket $\langle - \rangle$. We will also need the *replicated measure* $p_{\underline{X}^{(1)}|\underline{Y}}(\underline{x}^{(1)}|\underline{y}) p_{\underline{X}^{(2)}|\underline{Y}}(\underline{x}^{(2)}|\underline{y})$ on two

copies of the spin system and we use the same notation $\langle - \rangle$ for the associated bracket.

The special distributions of the quenched couplings (equal mean and variance) induce a gauge symmetry which implies a set of powerful identities due to Nishimori [7]. In particular if we define the magnetization and the overlap parameter

$$m_1 = \frac{1}{n} \sum_{i=1}^n x_i, \quad q_{12} = \frac{1}{n} \sum_{i=1}^n x_i^{(1)} x_i^{(2)}$$

we have the identity $P_{m_1}(x) = P_{q_{12}}(x)$ where

$$P_{m_1}(x) = \mathbb{E}_{\{h_i, J_c\}} \left[\left\langle \delta \left(x - \frac{1}{n} \sum_{i=1}^n x_i \right) \right\rangle \right]$$

$$P_{q_{12}}(x) = \mathbb{E}_{\{h_i, J_c\}} \left[\left\langle \delta \left(x - \frac{1}{n} \sum_{i=1}^n x_i^{(1)} x_i^{(2)} \right) \right\rangle \right]$$

Our simplification of the communication model consists in replacing the sparse Tanner graph by a complete hypergraph. The gauge symmetry is retained and this is fundamental for our analysis. The hamiltonian is then

$$\mathcal{H}_{p\text{-spin}}(\underline{x}) = - \sum_{i_1 < i_2 < \dots < i_p} J_{i_1, i_2, \dots, i_p} x_{i_1} x_{i_2} \dots x_{i_p} - \sum h_i x_i$$

where $J_{i_1, i_2, \dots, i_p} \sim \mathcal{N}(\frac{J p!}{2n^{p-1}}, \frac{J p!}{2n^{p-1}})$ and $h_i \sim \mathcal{N}(h, h)$. The mean and variance of J_{i_1, i_2, \dots, i_p} are normalized with n^{p-1} , to yield a non-trivial free energy, while the $\frac{p!}{2}$ is due to historical reasons.

Let us note that the standard p -spin model [?] has $h = 0$ and $J_{i_1, i_2, \dots, i_p} \sim \mathcal{N}(0, \frac{J p!}{2n^{p-1}})$ and hence no gauge symmetry. In the present model we retain this symmetry because it comes naturally from channel symmetry.

In the next section we explain our main result and its significance. The rest of the paper is devoted to a sketch of the analysis. Although here we limit ourselves to the AWGN channel, we can perform a similar analysis for any binary input memoryless symmetric channel. In the sequel we use the shorthand notation $\mathbb{E}[-]$ for all expectations with respect to quenched couplings (e.g channel outputs).

II. EXACT SOLUTION

The replica method applied to the p -spin model leads to a “Landau functional” (a function of the order parameter $m \in [0, 1]$),

$$f_{RS}(m) = \frac{J}{4} (1 - p m^{p-1} + (1-p) m^p) + \int \ln 2 \cosh(z \sqrt{J \frac{p}{2} m^{p-1} + h + J \frac{p}{2} m^{p-1} + h}) D z \quad (1)$$

The maximizer of this functional \hat{m} , is the equilibrium state magnetization and $f_{RS}(\hat{m})$ the free energy.

Our main result formulated below involves also another “functional”,

$$\tilde{f}(m) = \frac{J}{2} (1-p) m^p + \int \ln 2 \cosh(z \sqrt{J \frac{p}{2} \hat{m}^{p-1} + h + J \frac{p}{2} m^{p-1} + h}) D z \quad (2)$$

To our knowledge this expression does not have any direct physical significance and is related to the proof technique. We denote by \tilde{m} the maximizer of $\tilde{f}(m)$ and set

$$C_p = \{(J, h) | \hat{m} = \tilde{m}\}$$

It is not trivial to determine rigorously this region. We have proven that for $p = 2$ C_2 is equal to the whole two dimensional plane. For $p \geq 4$ C_p is definitely not equal to the whole plane but contains atleast $\{h \geq 0, J \in [0, K_1(p)] \cup [K_2(p), \infty)\}$ for some positive constants $K_{1,2}(p)$. Numerically we find a much bigger region for C_4 (see figure 1).

Our main result can now be stated

Theorem 1: Let p be even. For $(J, h) \in C_p$ the free energy and magnetization are given by the solution of the variational problem

$$\lim_{n \rightarrow \infty} \frac{1}{n} \mathbb{E}[\ln Z] = \max_{m \in [0, 1]} f_{RS}(m)$$

In the four remarks below we explain the content of this theorem.

Remark 1: We conjecture that the equality in the theorem should hold for any (J, h) .

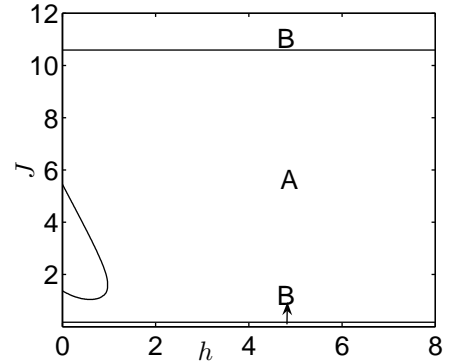


Fig. 1. The region C_4 in the space of J, h is numerically equal to $A \cup B$. Our rigorous estimates show that C_4 contains atleast B

Remark 2: As will be seen later the interpolation method yields that

$$\lim_{n \rightarrow \infty} \frac{1}{n} \mathbb{E}[\ln Z] \geq \max_{m \in [0, 1]} f_{RS}(m)$$

for all (J, h) . We stress that when the gauge symmetry is present (curiously) the interpolation method leads to the reverse bound compared to the standard Sherrington-Kirkpatrick or standard p -spin models.

Remark 3: The solutions \hat{m} of the variational problem are shown on figure 2. The left plot shows that the phase transition

is first order (jump discontinuity) for $p \geq 4$ and second order (continuous curve with jump in first derivative) for $p = 2$. The curves on the right show the magnetization as a function of the inverse noise h . In fact the curves are the cousins of the GEXIT functions of LDPC theory [8].

Remark 4: Since we have set $\mathbb{E}[h_i] = h \geq 0$ we get a non-negative magnetization. One may multiply h_i in the hamiltonian by $\eta = \pm 1$. Then the same results hold if $\eta = -1$ with $-\hat{m}$ replacing \hat{m} . Moreover below a critical value for J there are two coexisting Gibbs states (like in the usual Ising model) in the sense that $\lim_{h \rightarrow 0} \mathbb{E}[x_i] = \text{sign}(\eta)\hat{m}$.

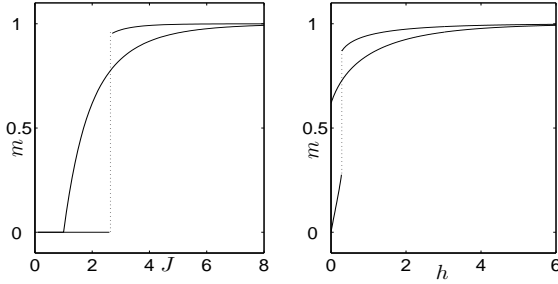


Fig. 2. Left: Plot of magnetization \hat{m} w.r.t J for $p = 2$ and 4 when $h = 0$. The magnetization jumps for $p = 4$ and is continuous for $p = 2$. Right: Plot of magnetization w.r.t h for $p = 2$ and 4 when $J = 2.5$

III. LOWER BOUND

The lower bound is based on the interpolation method developed by Guerra and Toninelli in [3]. We define a “mean field” Hamiltonian \mathcal{H}_{MF} where the spins do not interact with each other but instead experience an effective magnetic field. This “mean field” is random due to the randomness in the interaction coefficients of the original hamiltonian but the point is that its free energy is easy to evaluate. The basic idea of the interpolation method is to then define an appropriate interpolating Hamiltonian $\mathcal{H}_t(\underline{x})$ such that $t \in [0, 1]$ and $\mathcal{H}_{t=1} = \mathcal{H}_{p\text{-spin}}$ and $\mathcal{H}_{t=0} = \mathcal{H}_{MF}$.

The mean field hamiltonian is guessed by looking at the replica solution. The following is the hamiltonian whose free energy is given by the integral term (1)

$$\mathcal{H}_{MF}(\underline{x}) = - \sum_i J_i x_i - \sum_i h_i x_i$$

where $h_i \sim \mathcal{N}(h, h)$ and $J_i \sim \mathcal{N}(J \frac{pm^{p-1}}{2}, J \frac{pm^{p-1}}{2})$. The interpolating hamiltonian $\mathcal{H}_t(\underline{x})$ is simply

$$\mathcal{H}^{(t)}(\underline{x}) = \mathcal{H}_{p\text{-spin}}^{(t)}(\underline{x}) + \mathcal{H}_{MF}^{(t)}(\underline{x})$$

with couplings in $\mathcal{H}_{p\text{-spin}}^{(t)}$ distributed as $J_{i_1, i_2, \dots, i_p} \sim \mathcal{N}(\frac{tJp!}{2n^{p-1}}, \frac{tJp!}{2n^{p-1}})$, $h_i \sim \mathcal{N}(th, th)$ and in $\mathcal{H}_{MF}^{(t)}$ as $J_i \sim \mathcal{N}((1-t)J \frac{pm^{p-1}}{2}, (1-t)J \frac{pm^{p-1}}{2})$, $h_i \sim \mathcal{N}((1-t)h, (1-t)h)$. Notice that $\mathcal{H}^{(t)}(\underline{x})$ is carefully defined to retain the gauge symmetry.

Let us denote by $\alpha_n(t)$ and $\langle - \rangle_t$ the free energy and Gibbs average associated to $\mathcal{H}_t(\underline{x})$. We are interested in finding

$\frac{1}{n} \mathbb{E}[\ln Z] = \alpha_n(1)$ which can be written as

$$\frac{1}{n} \mathbb{E}[\ln Z] = \alpha_n(0) + \int_0^1 \frac{d\alpha_n(t)}{dt} dt$$

Using the integration by parts identity for $z \sim \mathcal{N}(0, \sigma^2)$, $\mathbb{E}[f(z)] = \sigma^2 \mathbb{E}[f'(z)]$ we find after some algebra

$$\begin{aligned} \frac{d\alpha_n(t)}{dt} &= \frac{J}{4} (1 - pm^{p-1} + (1-p)m^p) - \frac{J}{4} \mathbb{E}[\langle R(m, q_{12}) \rangle_t] \\ &\quad + \frac{J}{2} \mathbb{E}[\langle R(m, m_1) \rangle_t] + O\left(\frac{1}{n}\right) \end{aligned} \quad (3)$$

where

$$R(a, b) = (p-1)a^p - pa^{p-1}b + b^p$$

The Gibbs average $\langle - \rangle_t$ has gauge symmetry and hence $P_{m_1}^{(t)}(x) = P_{q_{12}}^{(t)}(x)$ which implies

$$\mathbb{E}[\langle R(m, q_{12}) \rangle_t] = \mathbb{E}[\langle R(m, m_1) \rangle_t]$$

Also,

$$\alpha_n(0) = \int \ln 2 \cosh(z \sqrt{J \frac{p}{2} m^{p-1} + h + J \frac{p}{2} m^{p-1} + h}) Dz$$

Thus we get the sum rule

$$\frac{1}{n} \mathbb{E}[\ln Z] = f_{RS}(m) + \int_0^1 \frac{J}{4} \mathbb{E}[\langle R(m, m_1) \rangle_t] dt + O\left(\frac{1}{n}\right) \quad (4)$$

Here (4) is true for all m and since $R(a, b)$ is non-negative for all a, b (for p even) we obtain the following bound

$$\lim_{n \rightarrow \infty} \frac{1}{n} \mathbb{E}[\ln Z] \geq \max_{m \in [0, 1]} f_{RS}(m) = f_{RS}(\hat{m})$$

IV. UPPER BOUND

The upper bound is obtained by showing that the remainder term in (4) goes to 0. For this purpose we define the following new hamiltonian

$$\mathcal{H}^{(t, \lambda)}(\underline{x}) = \mathcal{H}^{(t)}(\underline{x}) \Big|_{m=\hat{m}} - \frac{\lambda n}{2} R(\hat{m}, m_1)$$

If we can control the free energy $\alpha_n(t, \lambda)$ of this hamiltonian for small enough λ we will have an estimate of the remainder term in (4). This is because

$$\frac{\partial \alpha_n(t, \lambda)}{\partial \lambda} \Big|_{\lambda=0} = \frac{1}{2} \mathbb{E}[\langle R(\hat{m}, m_1) \rangle_t]$$

Differentiating along the path $\lambda(t) = \lambda_0 - J_0 t$ and using (3) we get after some algebra a sum rule for $\alpha_n(t, \lambda)$.

$$\begin{aligned} \alpha_n(t, \lambda(t)) &= \alpha_n(0, \lambda_0) + t \frac{J}{4} (1 - p\hat{m}^{p-1} + (1-p)\hat{m}^p) \\ &\quad - \frac{J}{4} \int_0^t \mathbb{E}[\langle R(\hat{m}, q_{12}) \rangle_{t, \lambda(t)}] dt + O\left(\frac{1}{n}\right) \end{aligned} \quad (5)$$

The convexity of $\alpha_n(t, \lambda)$ w.r.t λ implies if $\lambda(t) > 0$

$$\frac{\partial \alpha_n(t, \lambda)}{\partial \lambda} \Big|_{\lambda=0} \leq \frac{\alpha_n(t, \lambda(t)) - \alpha_n(t, 0)}{\lambda(t)}$$

The two sumrules (3) and (5) together with the positivity of $R(a, b)$ for even p assure us that the right hand side of this inequality is less than

$$\frac{\alpha_n(0, \lambda_0) - \alpha_n(0, 0)}{\lambda(t)} + \frac{1}{\lambda(t)} O\left(\frac{1}{n}\right)$$

Therefore we arrive at the following estimate for the remainder term

$$\mathbb{E} [\langle R(\hat{m}, m_1) \rangle_{t, \lambda(t)}] \leq \frac{1}{\lambda(t)} \left(\Delta_n(\lambda_0) + O\left(\frac{1}{n}\right) \right)$$

where

$$\Delta_n(\lambda_0) = \alpha_n(0, \lambda_0) - \alpha_n(0, 0)$$

It can be proven that if

$$\lim_{n \rightarrow \infty} \Delta_n(\lambda_0) = 0 \quad (6)$$

for all $\lambda_0 \leq J$ then

$$\lim_{n \rightarrow \infty} \int_0^1 \mathbb{E} [\langle R(\hat{m}, m_1) \rangle_t] dt = 0$$

We skip this argument here and sketch the proof of (6) in the next section. This will complete the proof of the theorem.

V. PROOF OF $\lim_{n \rightarrow \infty} \Delta_n(\lambda_0) = 0$

Notice that $\Delta_n(\lambda_0)$ is an increasing function of λ_0 . So, it is sufficient to show that $\lim_{n \rightarrow \infty} \Delta_n(J) = 0$. In fact we are able to evaluate exactly this limit by a sequence of saddle point calculations and it turns out that it vanishes if and only if $(J, h) \in C_p$. Let us denote $\lim_{n \rightarrow \infty} \Delta_n(J)$ as $\Delta(J)$.

A. The case $p = 2$

For the case of $p = 2$, there exists a simple way of evaluating $\Delta(J)$ using the approach taken in [3]. We obtain

$$\Delta(J) = \max_y g(y)$$

where

$$g(y) = \int Dz \ln \left[\frac{\cosh(y + \sqrt{\hat{m}J + h}z + J\hat{m} + h)}{\cosh(\sqrt{\hat{m}J + h}z + J\hat{m} + h)} \right] - \hat{m}y - \frac{y^2}{2J}$$

Case $J \leq 1$ and $h = 0$. We have $\hat{m} = 0$ and hence

$$\Delta(J) = \max_y g(y) = \max_y \left[\ln \cosh(y) - \frac{y^2}{2J} \right] = 0$$

Case $J \leq 1$ and $h > 0$. Equation $g'(y) = 0$ can have only one solution $y = 0$. Moreover, $g''(0) < 0$, therefore $g(y)$ is maximized at $y = 0$ and hence $\Delta(J) = g(0) = 0$.

Case $J > 1$ and $h = 0$. We have the following properties.

$$\begin{aligned} g(-y - J\hat{m}) &= g(y - J\hat{m}) \\ g'(y) &= 0 \text{ at } y = \{0, -\hat{m}J, -2\hat{m}J\} \\ g(0) &= g(-2\hat{m}J) = 0 \\ \lim_{y \rightarrow \infty} g(y) &= \lim_{y \rightarrow -\infty} g(y) = -\infty \end{aligned}$$

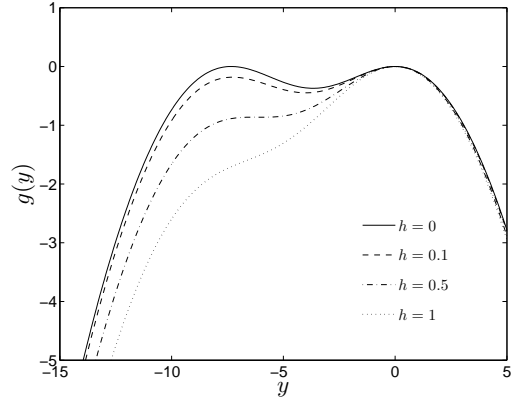


Fig. 3. Plot of $g(y)$ for various values of h , for fixed $J = 4$

Thus it is sufficient to show that $g(y)$ is maximum at $y = \{0, -2\hat{m}J\}$. But

$$g''(y) = \int \text{sech}^2(y + \sqrt{\hat{m}J}z + J\hat{m}) Dz - \frac{1}{J}$$

So if $g(0) = g(-2\hat{m}J)$ is not the maximum then,

$$g''(0) = g''(-2\hat{m}J) \geq 0 \text{ and } g''(-\hat{m}J) \leq 0$$

which would mean

$$\int \text{sech}^2(\sqrt{\hat{m}J}z) Dz \leq \int \text{sech}^2(\sqrt{\hat{m}J}z + J\hat{m}) Dz$$

a contradiction. Therefore $g''(0) = g''(-2\hat{m}J) \leq 0$ and $g''(-\hat{m}J) \geq 0$. Moreover $g(y)$ is maximized at $y = 0$ and $\Delta(J) = g(0) = 0$.

Case $J > 1$ and $h > 0$. A numerical evaluation indicates that $\Delta(J)$ is equal to 0 for all h_0 , but the above proof doesn't work. In fact the technique used for $p \geq 4$ works in this case also and allows to complete the proof.

The plot of the curve $g(y)$ is given in Figure 3 for $J = 4$. For very large negative y the dotted lines ($h > 0$) cross the full line ($h = 0$) so that it is not easy to reduce the $h > 0$ case to the $h = 0$ case.

B. General even p

For general even p , we cannot use the method of [3] to estimate $\Delta(J)$. Here we proceed differently using the following theorem.

Theorem 2: For a system with Hamiltonian

$$\mathcal{H}(\underline{x}) = - \sum J_i x_i - \frac{Jn}{2} m_1^p$$

where $J_i \sim \mathcal{N}(\mu, \sigma^2)$, the free energy is given by

$$\max_{m \in [0, 1]} \left[\frac{J}{2} (1-p)m^p + \int Dz \ln 2 \cosh(\sigma z + J \frac{p}{2} m^{p-1} + \mu) \right]$$

The proof which is based on saddle point calculations is quite lengthy and will be omitted here.

We notice that the condition $\Delta(J) = 0$ is equivalent to

$$\lim_{n \rightarrow \infty} \frac{1}{n} \mathbb{E} \left[\ln \sum_{\underline{x}} e^{\sqrt{\frac{p}{2} \hat{m}^{p-1}} J \sum J_i x_i + \sum h_i x_i + \frac{Jn}{2} m_1^p} \right] = \tilde{f}(\hat{m}) \quad (7)$$

Using theorem 2, with $\sigma = \sqrt{J \frac{p}{2} \hat{m}^{p-1} + h}$ and $\mu = h$, the left hand side of (7) can be written as $\max_m \tilde{f}(m)$, where we recall that $\tilde{f}(m)$ was defined in (2). Therefore, $\Delta(J) = 0$, if and only if $\max_m \tilde{f}(m) = \tilde{f}(\hat{m})$ which is true if and only if $\hat{m} = \tilde{m}$. This last condition is infact the definition of C_p .

VI. DETERMINATION OF THE REGION C_p

Below we sketch the analysis of the variational problems yielding \hat{m} , \tilde{m} and the domain C_p .

A. The case $p = 2$

The maxima of (1) and (2) are attained at one of their stationary points. For $p = 2$ these points are given by the solutions of the following fixed point equations respectively.

$$m = \int Dz \tanh(z\sqrt{Jm+h} + Jm+h) \quad (8)$$

$$m = \int Dz \tanh(z\sqrt{J\hat{m}+h} + Jm+h) \quad (9)$$

Case $h > 0$ and any J . Both (8),(9) have a unique positive solution which is the maximum of the respective free energies. So $\hat{m} = \tilde{m}$ for all J .

Case $h = 0$ and $J \leq 1$. Both (8) and (9) have only one solution (Figure 4) at $\hat{m} = \tilde{m} = 0$ and hence $\hat{m} = \tilde{m}$.

Case $h = 0$ and $J \geq 1$. Then (8) has two solutions $\{0, \hat{m}\}$ (Figure 4) and \hat{m} is the maximizer. Moreover (9) also has two solutions $\{0, \tilde{m}\}$, and \tilde{m} is the maximizer. Hence $\hat{m} = \tilde{m}$ for all $J > 0$

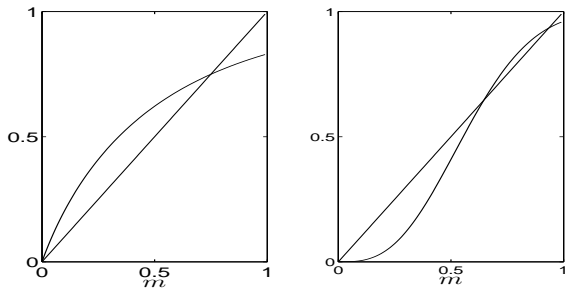


Fig. 4. Plot of (8) for $h = 0$. Left $p = 2$, right $p = 4$. Both the plots depict situations with different number of solutions to the fixed point equations

B. The case $p \geq 4$

For p even and $p \geq 4$, it is not always true that $\hat{m} = \tilde{m}$. This is because sometimes the corresponding fixed point equations

$$m = \int Dz \tanh(z\sqrt{J \frac{p}{2} m^{p-1} + h} + J \frac{p}{2} m^{p-1} + h) \quad (10)$$

$$m = \int Dz \tanh(z\sqrt{J \frac{p}{2} \hat{m}^{p-1} + h} + J \frac{p}{2} m^{p-1} + h) \quad (11)$$

have 3 solutions with 2 of them being local maxima (figure 4). So \hat{m} need not be equal to \tilde{m} and hence there are values $(J, h) \notin C_p$.

Case of small J . If $J \leq \frac{2}{p(p-1)} \equiv K_1(p)$, (10) and (11) have only one solution which maximize (1) and (2) respectively. So, $\hat{m} = \tilde{m}$ and $(J, h) \in C_p$ for $J \leq K_1(p)$.

Case of large J and $h = 0$. Then (1) has one more local maximum at $m \approx 1$ in addition to the one at 0. (e.g, when $p = 4$, $\hat{m} \approx 0.992$ for $J = 4$). We note that

$$f_{RS}(1) > \int Dz \left(\sqrt{J \frac{p}{2}} z + J \frac{p}{2} \right) + \frac{J}{4}(2-2p) = \frac{J}{2}$$

$$f_{RS}(0) = \ln 2 + \frac{J}{4}$$

Therefore, for sufficiently large J ($\approx 4 \ln 2$), \hat{m} is equal to the solution of (10) close to 1. Now $\tilde{f}(m)$ also has two local maximums at 0 and $m \approx 1$ and

$$\tilde{f}(1) > \int Dz \left(\beta \sqrt{J \frac{p}{2}} z + J \frac{p}{2} \right) + \frac{J}{2}(1-p) = \frac{J}{2}$$

$$\tilde{f}(0) = \int Dz \ln 2 \cosh \left(\sqrt{J \frac{p}{2}} z \right) < \ln 2 + \sqrt{\frac{Jp}{2}}$$

Therefore, if $J \geq p + 2 \ln 2 + \sqrt{p(p+4 \ln 2)} \equiv K_2(p)$, $f_{RS}(0) < f_{RS}(1)$ and $\tilde{f}(0) < \tilde{f}(1)$, hence the constraint $\hat{m} = \tilde{m}$ is satisfied. Summarising, if $h = 0$, $(J, h) \in C_p$ for $J \geq K_2(p)$.

Case of large J and $h > 0$. The maxima for both the (1),(2) (i.e, \hat{m}, \tilde{m}) will favour the stationary point close to $m = 1$ and hence for $h \geq 0$ and $J \geq K_2(p)$ we have $\hat{m} = \tilde{m}$.

Summarizing we find that C_p is atleast as big as $\{h \geq 0, J \in [0, K_1(p)] \cup [K_2(p), \infty)\}$.

ACKNOWLEDGMENT

The work presented in this paper was partially supported by the National Competence Center in Research on Mobile Information and Communication Systems (NCCR-MICS), a center supported by the Swiss National Science Foundation under grant number 5005-67322.

REFERENCES

- [1] N. Sourlas, "Spin glass models as error correcting codes", *Nature*, **339** pp. 693-695, 1989.
- [2] A. Montanari, "The glassy phase of Gallager codes," *European Physical Journal*, **23** (2001)
- [3] F. Guerra and F.L. Toninelli, "Quadratic replica coupling in the Sherrington-Kirkpatrick mean field spin glass model", *J. of Math. Phys.*, **43**, Issue 7, p.3704, "Oct", 2002
- [4] S. Franz, M. Leone, "replica bounds for optimization problems and diluted spin systems", *Journal of Statistical Physics* **111**, (2003) 535-564
- [5] A. Montanari, "Tight Bounds for LDPC and LDGM codes," *IEEE Transactions on Information Theory*, **51**, Issue 9, (2005) pp:3221 - 3246
- [6] M. Talagrand, "The Parisi Formula", *Annals of Mathematics* **163**, Issue. 1, pp: 221-263
- [7] H. Nishimori, "Statistical Physics of Spin Glasses and Information Processing", *Oxford Science Publications*,
- [8] T. Richardson, R. Urbanke, "Modern coding theory", book in preparation